第28卷 第3期　　工　程　数　学　学　报　　Vol. 28 No. 3
2011年06月　　CHINESE JOURNAL OF ENGINEERING MATHEMATICS　　June 2011

文章编号:1005-3085(2011)03-0401-05

股票熵风险度量方法研究*

袁　博[1,2]，　王建国[2]

(1- 河北北方学院研究生部，张家口　075000;　2- 西安建筑科技大学理学院，西安　710055)

摘　要: 本文将熵理论引入到股票投资领域，提出了适用于中国股票市场的熵风险度量方法．该方法将股票收益区间均分成若干个子区间，利用收益率落在各收益子区间中的频数度量熵风险值，并引入调节因子将初始熵风险值标准化．实证算例选取上海证券交易所股票数据对该度量方法进行检验．检验结果表明，初始熵风险值随收益子区间数目指数的增大而递增，标准化后的熵风险值在收益子区间数目指数取到5后趋于稳定，说明熵风险度量方法能有效评估股票风险．

关键词: 熵；风险度量；收益子区间

分类号: **AMS**(2000) 90A09　　**中图分类号**: F830　　**文献标识码**: A

1　引言

美国经济学家Markwitz提出用方差度量投资风险，开创了投资风险度量的量化时代[1]．自此之后，各种投资风险度量方法层出不穷．被广泛认可的方法包括：方差度量方法、系数度量方法等等．但是，这些度量方法又各自存在着的弊端，比如方差方法受收益分布必须服从正态分布的假设以及系数度量方法对单只股票风险与股市总收益满足线性关系的假设条件通常很难实现[2]．

由于“熵”主要是用来度量信息的无序性或不确定性的，因此，也可将“熵”引入到金融投资领域，度量投资的风险．研究表明，作为风险度量手段，熵方法体现了概率分布在整个取值空间的平均不确定程度，并且反映了损失分布的大小．尽管如此，熵度量投资风险方法对时间动态变化不敏感，而且熵风险计算公式中所要求的收益状态概率也较难确定．针对这一问题，本文提出了通过计算收益率落在收益子区间内的频数求度量熵风险值的可实际操作方法，并论证该方法的合理性，进而给出了实现策略；最后，实证算例检验本方法的有效性．

2　熵风险度量方法

由于证券市场受经济政策和资金流动等不确定因素的影响，投资者的实际收益往往与决策时期望的收益不一致，即投资存在风险．在投资学中，投资风险可被刻画为未来收益的不确定性及其发生的概率[3]．

假设状态 $x_i\,(i=1,2,\cdots,n)$ 是以概率

$$p(x_i),\quad i=1,2,\cdots,n,\quad 0\le p(x_i)\le 1,\quad \sum_{i=1}^{n}p(x_i)=1$$

收稿日期: 2009-05-04.　**作者简介**: 袁博 (1978年7月生)，男，讲师. 研究方向：金融数学.

***基金项目**: 国家自然科学基金 (60974140)；张家口市科技项目指导计划 (1021002B).

出现的随机事件，状态x_i的不确定性程度可由函数$L(x_i) = -\ln p(x_i)$予以度量．若投资方案的状态空间$X := \{x_1, x_2, \cdots, x_n\}$中状态$x_i$是以概率$p(x_i)$出现的随机事件，那么可以用状态$x_i$的不确定性程度来度量风险的大小，即可用函数$L(x_i)$表示风险函数，它是概率为$p(x_i)$的状态$x_i$的风险测度函数．显然，$L(x_i)$的量值由状态$x_i$唯一决定，是状态空间为$X$: $\{x_1, x_2, \cdots, x_n\}$，概率空间为$P$: $\{p_1, p_2, \cdots, p_n\}$的不确定程度的测度函数关于状态$x_i$的分量．因而，熵函数

$$H(p_1, p_2, \cdots, p_n) = -\sum_{i=1}^{n} p(x_i) \ln p(x_i)$$

可看作是状态空间为X，概率空间为P的投资方案的风险度量．熵函数$H(p_1, p_2, \cdots, p_n)$从期望的角度度量投资行动方案中所有状态具有的风险，它是投资方案总体风险的测度函数，由投资行动方案概率空间P总体结构唯一决定．该方法称为熵风险度量方法．

与已有的风险度量方法相比较，熵风险度量方法有以下几个优点：

1) 熵的概念来源于热力学和信息学，它能够很好地反映微观系统的无序程度或不确定程度．股票投资的风险来源于收益状态的波动性，收益状态的波动也是复杂和无序的．因此，利用熵方法可更客观地反映投资风险．而常规的风险度量方法中，方差方法、VaR 度量方法、β系数度量方法等对收益状态波动性的无序程度把握不够，很难达到熵方法对投资风险度量的有效程度．在熵函数$H(p_1, p_2, \cdots, p_n)$中，所有状态$x_1, x_2, \cdots, x_n$的波动对收益的影响都予以充分地考虑[4]；

2) 在现有的风险度量方法中，需要投资者对投资收益做一个预期，以实际收益对预期收益的偏差作为风险的测度函数，例如方差度量方法．这样的度量方法存在着明显的缺点：首先，对投资收益做预测，人为的主观因素会影响风险度量的精确性．其次，这一类度量方法主要是刻画实际收益对预期收益偏差的度量，这种对应属于单点对应，并且是一阶矩估计，很难体现出概率分布在整个空间上的不确定程度，反映出的分布信息明显不足．而熵风险度量函数是通过收益波动度量投资风险，即从客观角度度量投资风险，不受人为因素影响；

3) 现有的风险度量方法中，要求的条件较为苛刻．譬如，诸多度量方法要求股票投资的收益分布满足正态分布或对称的近似正态分布．但事实上，当股票投资收益受某一特定因素影响波动性非常大时，对股票收益服从正态分布或者是对称分布的假设非常牵强，即使样本容量很大也很难实现近似的正态分布．现实证券市场中，股票收益大多呈现“尖峰厚尾”状态，在这种情况下，如果使用 VaR 度量方法就会产生较大的度量误差[5,6]．根据资本资产定价理论的假设，对于β系数度量方法，要求单支股票的收益与股票市场收益满足线性关系．但实际情况是，市场并不是决定股票价格的唯一因素，政策、局势甚至某一特定的消息都会引起股票价格的巨大波动，这使单支股票收益与股票市场总收益满足线性关系的假设很难实现．而熵风险度量方法对于投资环境的要求很低，只要是有效市场理论支持的环境，熵风险度量方法都适用．在熵风险理论中，收益的波动源于信息的不确定性．因此，熵能够准确反映收益的波动，作为风险度量手段也更合理[7]．

下面我们给出基于熵函数的风险度量$H(p_1, p_2, \cdots, p_n)$的具体计算方法．

根据我国证券管理规定，正常情况下股票在单日内涨跌幅度被限定在10%以内，即股票实际收益区间为$[-10\%, 10\%]$，将收益区间均分为q个收益子区间

$$[-10\%, 10\%] = \bigcup_{i=1}^{q} \Delta_i, \quad i = 1, 2, \cdots, q. \tag{1}$$

为了研究子区间数对熵风险度量值的影响，q可取值如下

$$q = 10 \times 2^k, \tag{2}$$

其中q为子区间数，k定义为子区间数目指数．由(2)式，当子区间数目指数分别取$k = 0,1,2,\cdots,9$时，对应的子区间数分别为$q = 10,20,40,80,160,320,640,1280,2560,5120$.

设$r_j\,(j = 1,2,\cdots,d)$为第d支股票的日收盘价，d为数据采集天数，

$$R_j = \frac{r_j - r_{j-1}}{r_{j-1}}$$

为日收益率，$-10\% \le R_j \le 10\%$．将区间$[-10\%,10\%]$均分成$q = 10 \times 2^k$个子区间，取步长为$l = \frac{0.2}{q}$，子区间表示为

$$\triangle_i = \begin{cases} [-10\% + (i-1)l, -10\% + il\,), & i = 1,2,\cdots,q-1, \\ [-10\% + (i-1)l, -10\% + il\,], & i = q. \end{cases} \tag{3}$$

假设R_j落在第i个子区间内的次数记为n_i，记频数$n_i = \frac{n_i}{d}$，则定义

$$H_p(k) = -\sum_{i=1}^{q} \rho_i \ln \rho_i \tag{4}$$

为第p支股票的初始熵风险值.

假设有m支股票，将子区间数目指数为k的m支股票的初始熵值的均值定义为初始熵均值，记为

$$\overline{H}_k = \frac{1}{m} \sum_{p=1}^{m} H_p(k). \tag{5}$$

通过验证发现，本文实证部分的初始熵均值$\overline{H}$与子区间数目指数k具有良好的线性关系，因此可以选取线性的调节因子$A + Bk$，使求得的初始熵风险值被限定在一个固定范围内，得到标准化后的熵风险值，定义

$$H^* = -\sum_{i=1}^{q} \rho_i \frac{\ln \rho_i}{A + Bk} \tag{6}$$

为股票标准熵风险值.

初始熵均值是对初始熵风险值求均，初始熵均值与子区间数目指数具有良好的线性关系，据此引入线性调节因子．初始熵风险值与标准熵风险值是在调节因子取不同值情况下的熵风险值．当调节因子为1时，标准熵风险值即为初始熵风险值．初始熵风险值会随着子区间数目指数的增加而增大，取线性调节因子$A + Bk$时，标准熵风险值趋向于常数.

3 实证研究

从上海证券交易所选取50支来自于不同板块、行业的股票$S_1, S_2, \cdots, S_{50}$如下：

600000浦发银行、600005武钢股份、600009上海机场、600019宝钢股份、600028中国石化、600030中信证券、600066宇通客车、600085同仁堂、600098广州控股、600350山东高

速、600138中青旅、600150中国船舶、600169太原重工、600176中国玻纤、600219南山铝业、600238海南椰岛、600295鄂尔多斯、600323南海发展、600997开滦股份、600251冠农股份、600351亚宝药业、600362江西铜业、600391成发科技、600406国电南瑞、600479千金药业、600495晋西车轴、600511国药股份、600519贵州茅台、600525长园新材、600547山东黄金、600588用友软件、600600青岛啤酒、600612中国铅笔、600663陆家嘴、600685广船国际、600690青岛海尔、600702沱牌曲酒、600717天津港、600729重庆百货、600737中粮屯河、600798宁波海运、600809山西汾酒、600820隧道股份、600822上海物贸、600858银座股份、600879火箭股份、600880博瑞传播、600887伊利股份、600897厦门空港、600900长江电力.

选取了自2005年元月4日起连续四年的日收盘价数据，取 $d=996$ 进行实验.

首先取调节因子 $A+Bk=1$，利用Matlab 7.01编程计算50支股票的初始熵风险值，并对初始熵值求均值，表1列出了 $S_{10},S_{20},\cdots,S_{50}$ 的初始熵风险值与初始熵均值.

表1: 初始熵值表

股票代码	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$
S_{10}	1.688	2.345	3.020	3.676	4.314	4.879	5.348	6.210	7.010	7.488
S_{20}	2.031	2.692	3.323	3.955	4.588	5.132	5.467	6.239	7.097	7.476
S_{30}	2.048	2.717	3.366	3.999	4.610	5.177	5.712	6.421	7.350	7.665
S_{40}	1.984	2.608	3.219	3.861	4.465	5.038	5.584	6.554	6.885	7.446
S_{50}	1.479	2.132	2.800	3.461	4.120	4.715	5.264	6.299	7.016	7.613
初熵均值 $\overline{H}_k$	1.849	2.513	3.174	3.824	4.455	5.040	5.690	6.236	7.064	7.647

由表1可以看出，初始熵风险值会随着子区间数目指数的增加而增长. 而且初始熵均值 $\overline{H}_k$ 与子区间数目指数 k 之间存在较强的线性关系. 将初始熵均值 $\overline{H}_k$ 与 k 线性拟合，则有调节因子为 $0.63695k+1.87638$，代入(6)式，得

$$H^*=-\sum_{i=1}^{q}\rho_i\frac{\ln\rho_i}{0.63695k+1.87638}$$

为标准熵风险值. 应用Matlab软件，计算50支股票标准熵风险值，仍然只列出 $S_{10},S_{20},\cdots,S_{50}$ 的标准熵风险值，结果见表2.

表2: 标准熵风险值表

股票代码	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$
S_{10}	0.672	0.728	0.769	0.793	0.807	0.806	0.806	0.806	0.806	0.806
S_{20}	0.808	0.836	0.846	0.853	0.859	0.848	0.846	0.848	0.847	0.842
S_{30}	0.815	0.843	0.857	0.863	0.863	0.855	0.856	0.857	0.857	0.856
S_{40}	0.789	0.809	0.819	0.833	0.836	0.833	0.832	0.832	0.833	0.832
S_{50}	0.588	0.662	0.713	0.745	0.771	0.779	0.781	0.781	0.781	0.780

由表2可以看出，当收益子区间数目指数增加步长减小时，度量结果对子区间数目指数为5的标准熵风险值趋近效果较好. 通过计算，在收益子区间数目指数由0增加至5的过程中，股票的标准熵风险值逐渐趋向于常数值0.85，该常数为对股票风险度量的标准熵风险值. 标准熵风险值为0.85表明中国的股票市场风险较大.

4 结论

本文利用熵方法度量股票风险，将收益区间均分为若干收益子区间，通过计算收益率落在每个子区间的频数求初始熵风险值，再引入调节因子将初始熵风险值标准化．实证研究表明标准熵风险值在子区间数目指数取5后达到稳定．本文中股票熵风险度量方法具备以下特点：1) 在评估股票业绩的诸多数据中，股票的日收益率最基础、最稳定，而熵风险度量方法数据源恰好取自日收益率，这确保了熵风险度量方法的实用性；2) 本文中熵风险度量方法是从股票收益稳定性评估角度考虑的，而在信息论中熵是与已知信息构成互补的因素，是不确定性最好的计量工具，因此利用熵度量股票风险又极具合理性；3) 熵风险度量法无需讨论分布对称性，收益分布更不要求是正态的，充分体现出熵风险度量方法的高效性．

参考文献：

[1] Ravhev S T, Frank, Faboch. The proper use of measures in portfolio theory[J]. International Journal of Theoretical and Applied Finance, 2008, 8(8): 1107-1133

[2] 李华，李兴斯．均值-叉熵股票投资组合优化模型[J]．数学的实践与认识，2005, 35(5): 65-70
Li H, Li X S. Mean-cross-entropy model for portfolio optimization[J]. Mathematics in Practice and Theory, 2005, 35(5): 65-70

[3] Nguyen Q B, Mebarki A, *el al.* Integrated probabilistic framework for domino effect and risk analysis[J]. Advances in Engineering Software, 2009, 40: 892-901

[4] Fischer G, Ermoliva T. Livestock production planning under environmental risks and uncertainties[J]. Journal of Systems Science and Systems Engineering, 2006, 15(4): 399-418

[5] Benth F E, Meyer B. The density process of the minimal entropy martingale measure in a stochastic volatility model with jumps[J]. Finance Stock, 2008, 9(4): 563-575

[6] Seiford L M, Zhu J. Modeling undesirable factors in efficiency evaluation[J]. European Journal of Operational Research, 2009, 142: 16-20

[7] Takeda A, Takafumi K. A robust approach based on conditional value-at-risk measure to statistical learning problems[J]. European Journal of Operational Research, 2009, 198: 287-296

Study on the Entropy-based Risk Measure Method

YUAN Bo[1,2], WANG Jian-guo[2]

(1- Department of Graduate, Hebei North University, Zhangjiakou 075000;
2- College of Science, Xi'an University of Architecture and Technology, Xi'an 710055)

Abstract: In this paper, the entropy theory is introduced to the stock investment field, based on which a risk measure method is proposed. In the method, the investment bonus interval is partitioned into a number of isometric subintervals and the entropy-based risk value is determined as the frequency of stock bonus that falls in each of subintervals. Further, a kind of tuning factor is introduced to standardize the initial entropy-based risk value. The validity and effectiveness of the proposed method is testified by a group of empirical stock data of Shanghai Exchange Institution. It shows that the *initial entropy-based risk value increases with the number of the subintervals.* It also shows that the entropy-based risk value is stabilized at a constant while the number of subinterval reaches to 320. The empirical result demonstrates that the proposed entropy-based risk measure method is practical to assess the investment risk.

Keywords: entropy; risk measure; bonus subinterval

Received: *04 May 2009.* **Accepted:** *21 Sep 2010.*
Foundation item: The National Natural Science Foundation of China (60974140); the Science and Technology Guide Project of Zhangjiakou City (1021002B).